Sampling Distribution of $\overline{x}$

AP Statistics · Mr. Merrick · January 16, 2026

Suppose we take a random sample of size n from a population.

- Let $X_1, X_2, \ldots, X_n$ be the sampled values.
- Let μ be the true population mean, and σ the true population standard deviation.
- For a finite population of size N with values $v_1, \ldots, v_N$,

$$\mu = \frac{1}{N}\sum_{i=1}^{N} v_i \qquad \text{and} \qquad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(v_i - \mu)^2}.$$

- The sample mean is

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Because the sample is random, $\overline{x}$ is a **random variable**. To make probability statements about $\overline{x}$, we must decide whether a **normal model** is appropriate. That depends on:

- whether observations can be treated as **independent**, and
- whether the sampling distribution of $\overline{x}$ is **(approximately) normal**.

Normality for $\overline{x}$ can come from either:

- the **population is (approximately) normal** *(then $\overline{x}$ is normal for any n)*, or
- the **Central Limit Theorem (CLT)** *(if $n \geq 30$, then $\overline{x}$ is approximately normal).*

	Independence OK ($n \leq 0.10N$)	**Independence NOT OK** ($n > 0.10N$)
Normality NOT OK (*population not normal & $n < 30$*)	**No reliable normal model** $\overline{x}$ may be skewed. Use simulation/technology or reconsider conditions. (For inference, consider a larger n.)	**No reliable normal model (finite population)** Dependence and non-normality both interfere. Exact/technology methods may be needed (simulation or enumeration for small N).
Normality OK (*population normal OR $n \geq 30$*)	**Normal model for $\overline{x}$** If normality is reasonable and independence holds, then $$\overline{x} \approx \mathcal{N}\left(\mu,\ \frac{\sigma}{\sqrt{n}}\right).$$	**Normal model with finite population correction** When sampling without replacement and the 10% condition fails, $$\overline{x} \approx \mathcal{N}\left(\mu,\ \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right).$$ This is the **FPC** adjustment.

Important inference note (AP Stats): In practice, we usually do *not* know σ. When normality is reasonable, we estimate σ with the sample standard deviation s and use a **t-model for inference about** μ. (This handout focuses on the *sampling distribution* of $\overline{x}$.)

1. For each situation below, identify the **most appropriate model/approach** for the sampling distribution of $\overline{x}$. *No probability calculations are required. Justify your choice.*

   (a) A nutritionist takes a random sample of $n = 12$ teenagers from a very large city and records daily sugar intake (grams). The population distribution of sugar intake is strongly right-skewed.

   (b) A factory claims the mean diameter of ball bearings is $\mu = 8.00$ mm with population standard deviation $\sigma = 0.04$ mm. A random sample of $n = 36$ bearings is taken from a day's production (huge production run).

   (c) A teacher randomly selects $n = 20$ tests *without replacement* from a stack of $N = 120$ tests to check the average score. The distribution of scores in the stack is approximately normal with known μ and σ.

   (d) A researcher samples $n = 25$ people from a population known to be normally distributed for systolic blood pressure.

   (e) A small company has $N = 40$ employees. A manager selects $n = 18$ employees without replacement and records commute times. The employee commute-time distribution is strongly right-skewed.

2. A brand of granola bars has mean fat content $\mu = 7.5$ grams and population standard deviation $\sigma = 1.2$ grams. A random sample of $n = 64$ bars is selected from a very large shipment.

   (a) Identify the appropriate model for the sampling distribution of $\overline{x}$.

   (b) Find $P(\overline{x} > 7.8)$.

3. A town has $N = 500$ households. The mean household water use last month was $\mu = 12.4$ (thousand gallons) with population standard deviation $\sigma = 3.0$ (thousand gallons). A simple random sample of $n = 80$ households is taken without replacement.

   (a) Explain why the usual independence condition fails.

   (b) If the population distribution of water use is roughly normal, find the approximate probability that $\overline{x} < 11.8$. Use the finite population correction.

4. A college reports that the mean time students spend studying per week is $\mu = 14$ hours with population standard deviation $\sigma = 6$ hours. A random sample of $n = 30$ students is selected.

   (a) Explain why a normal model for $\overline{x}$ is reasonable even if the population distribution is somewhat skewed.

   (b) Find $P(13 \leq \overline{x} \leq 16)$.

5. A hospital states that the mean waiting time in the ER is $\mu = 42$ minutes. Assume the population standard deviation is $\sigma = 18$ minutes. A random sample of $n = 50$ patients is selected from a very large time period.

   Determine the probability that the sample mean waiting time is less than 38 minutes. Clearly justify your model choice.